



Bennie, S. J., Curchod, B. F. E., Manby, F. R., & Glowacki, D. R. (2017). Pushing the Limits of EOM-CCSD with Projector-Based Embedding for Excitation Energies. *Journal of Physical Chemistry Letters*, 8(22), 5559-5565. <https://doi.org/10.1021/acs.jpcllett.7b02500>

Publisher's PDF, also known as Version of record

License (if available):  
CC BY

Link to published version (if available):  
[10.1021/acs.jpcllett.7b02500](https://doi.org/10.1021/acs.jpcllett.7b02500)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

This is the final published version of the article (version of record). It first appeared online via American Chemical Society at <https://doi.org/10.1021/acs.jpcllett.7b02500>. Please refer to any applicable terms of use of the publisher.

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# Pushing the Limits of EOM-CCSD with Projector-Based Embedding for Excitation Energies

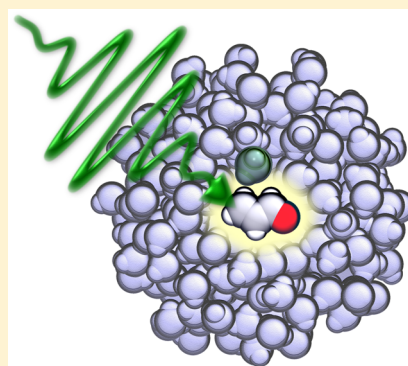
Simon J. Bennie,<sup>\*,†</sup> Basile F. E. Curchod,<sup>\*,†</sup> Frederick R. Manby,<sup>\*,†</sup> and David R. Glowacki<sup>\*,†,‡</sup>

<sup>†</sup>Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol, BS8 1TH, United Kingdom

<sup>‡</sup>Department of Computer Science, University of Bristol, Bristol, BS8 1UB, United Kingdom

## S Supporting Information

**ABSTRACT:** The calculation of accurate excitation energies using *ab initio* electronic structure methods such as standard equation of motion coupled cluster singles and doubles (EOM-CCSD) has been cost prohibitive for large systems. In this work, we use a simple projector-based embedding scheme to calculate the EOM-CCSD excitation energies of acrolein solvated in water molecules modeled using density functional theory (DFT). We demonstrate the accuracy of this approach gives excitation energies within 0.01 eV of full EOM-CCSD, but with significantly reduced computational cost. This approach is also shown to be relatively invariant to the choice of functional used in the environment and allows for the description of systems with large numbers of basis functions (>1000) to be treated using state-of-the-art wave function methods. The flexibility of embedding to select orbitals to add to the excited-state method provides insights into the origins of the excitations and can reduce artifacts that could arise in traditional linear response time-dependent DFT (LR-TDDFT).



The extension of density functional theory to the time-dependent domain (TDDFT) by Runge and Gross<sup>1,2</sup> and its efficient linear response implementation (LR-TDDFT) by Casida<sup>3</sup> has led to the practical calculation of the excited-state properties for medium and large molecular systems of hundreds of atoms.<sup>4–9</sup> LR-TDDFT generally is used when system size becomes a consideration due to its low scaling, but comes with the caveat that practical implementations, which employ an approximate exchange and correlation functional and the adiabatic approximation,<sup>4,5,7</sup> do not have reliable accuracy for charge-transfer excitations,<sup>10,11</sup> transitions with double-excitation character,<sup>12,13</sup> or  $S_1/S_0$  conical intersections.<sup>14</sup> From a practical standpoint, the main disadvantage of LR-TDDFT arises from the wide choice of functionals, each of which may have certain failure cases depending on the systems and physical process under study. For high accuracy benchmark calculations of excited-state properties, *ab initio* correlation methods are needed, although traditionally the size of the systems that can be treated by these methods is limited to 10–20 heavy atoms in a moderate basis. The gold-standard method is equation of motion coupled cluster singles and doubles (EOM-CCSD),<sup>15,16</sup> which is more or less black-box and can handle a wide range of chemical problems with predictable accuracy. However, EOM-CCSD has a worse formal scaling compared to LR-TDDFT.

Various approaches have been developed to improve the scaling and cost of EOM-CCSD, such as local orbital methods,<sup>17–21</sup> restricted virtual spaces and highly efficient implementations.<sup>22,23</sup> For larger systems, however, methods with such high scaling often require a multiscale approach such

as the quantum mechanical molecular mechanical (QM/MM) type partitioning<sup>24–27</sup> or other multilevel techniques.<sup>28–30</sup> The aim of multiscale modeling is to treat different regions of a system with methods matched to the required accuracy, e.g., important regions are treated with higher accuracy methods, and less important regions with lower accuracy methods. The accuracy/cost trade-off of such an approach is often better than treating the entire system with the same method. It is in this spirit that we present work that extends projector-based embedding<sup>31–34</sup> to excited-state properties. We present an approach to conduct density-functional/Hartree–Fock calculations on a large system, select a chemical subregion of interest, and embed a more expensive method inside it. This avoids the expense that would be incurred by applying the high-level method to the whole system. The high-level method becomes polarized by the low-level method, allowing us to obtain very accurate results. For excited-state calculations, the only extra knowledge needed for this approach concerns the atoms most involved in the excitation, or for improved efficiency the exact orbitals that may be needed.

We shall show that it is possible to obtain accurate excited-state excitation energies and oscillator strengths from embedded EOM-CCSD (eEOM-CCSD). We have selected one small test system of a water and formaldehyde to highlight physical effects of embedding, followed by a larger acrolein in water the efficiency and accuracy of this method. We have

**Received:** September 21, 2017

**Accepted:** October 27, 2017

**Published:** October 27, 2017



focused on a noncovalent system for embedding to show it is possible to obtain solvent effects cheaply and avoid artifacts that are known to occur in such systems;<sup>35</sup> however, we also show that it is possible to partition covalently bound systems by selecting important orbitals on the water to improve accuracy. It should be stressed that there is no fundamental reason that this approach would not be appropriate for obtaining excited-state properties for large covalent molecules as long as the excitation can be localized.

**Projector-Based Embedding for Excited States.** The detailed description of the embedding method has been presented elsewhere.<sup>31–34</sup> Here we briefly reiterate the main equations and features of projector-based embedding to keep this paper self-contained. Projector-based embedding belongs to a class of quantum embedding techniques that constructs a low-level potential from a method such as DFT and uses it to polarize a higher-level method such as CCSD(T). This type of embedding partitions the electronic density matrix into subregions A and B

$$\rho^{\text{Total}} = \rho^{\text{A}} + \rho^{\text{B}} \quad (1)$$

In the case of projector embedding, the partitioning is constructed by first localizing the molecular orbitals, and then automatically categorizing whether they are primarily on atoms in the active region ( $\rho^{\text{A}}$ ) or the environment ( $\rho^{\text{B}}$ ) based on a population analysis. In this work we use the standard Mulliken analysis to select our molecular orbitals (MOs) based on if the majority of the density (by default >0.4 electrons) resides on basis functions centered on atoms in the active region. We will highlight in our results that for excited-state properties a better selection criterion may be possible based on the configuration interaction (CI) vector.

The partitioning of the electronic density based on the Kohn–Sham orbitals<sup>36</sup> makes it possible to define the potential for the subregions free of nonadditive errors. The Fock operator for subsystem A can be built from the partitioned reduced one particle density matrix  $\gamma$  for the active region (A) and environment (B)

$$\mathbf{F}^{\text{A}} = \mathbf{H} + \mathbf{J}[\gamma^{\text{A}} + \gamma^{\text{B}}] + \mathbf{v}_{\text{xc}}[\gamma^{\text{A}} + \gamma^{\text{B}}] + \mu\mathbf{P}^{\text{B}} \quad (2)$$

where the first term is the normal core Hamiltonian, the second is the Coulomb potential, the third term is the exchange–correlation potential, and the final term is the level shift operator. This last term has the effect of enforcing the subsystems to be orthogonal, ensuring that the Pauli principle is obeyed and restricting the wave function of the active region to not over-relax into the environment (subsystem B). The  $\mu$  parameter is a large numerical value ( $10^6$ – $10^{10}$ ) that elevates all of the orbitals that construct  $\gamma^{\text{B}}$  to an extremely high energy, effectively making them inaccessible to the active region orbitals that build  $\gamma^{\text{A}}$ .  $\mathbf{P}^{\text{B}}$  is constructed out of the subsystem B (environment) density and the overlap matrix  $\mathbf{S}$ ,

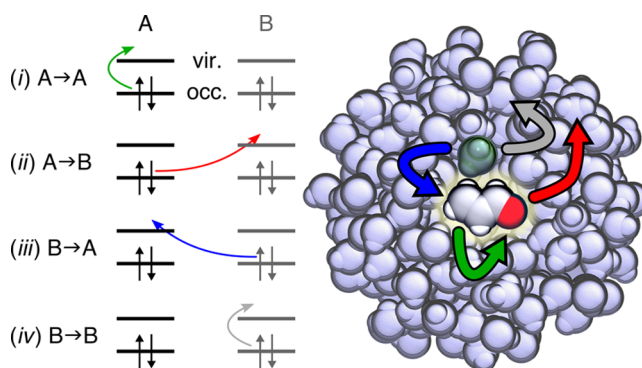
$$\mathbf{P}^{\text{B}} = [\mathbf{S}\gamma^{\text{B}}\mathbf{S}] \quad (3)$$

A key feature of projector embedding is that it exactly reconstructs the environment potential, in the sense that the energy (and other properties) of DFT-in-DFT is the same as DFT on the whole system. The method is simple to implement, relying only on a modified core Hamiltonian. This allows most correlation methods (e.g., MP2,<sup>37</sup> CCSD(T),<sup>38</sup> MRCI) to be performed in the high-level region without bespoke reprogramming. It is also simple to accelerate by using basis-set truncation

to further reduce the cost of the correlation method by reducing the scaling with the number of virtual orbitals.<sup>33</sup> We note here that the frozen-density embedding strategy was combined with LR-TDDFT for excitation energies.<sup>39,40</sup>

One potential concern of using projector-based embedding for excited-states is that occupied orbitals in the environment (subsystem B) will not be able to contribute to the excited-state relaxation as they are frozen in the form of the ground state and only exist implicitly. The simplest way to improve the description of subsystem A is to include extra atoms in subsystem A; however, a more efficient and instructive approach is to determine those occupied orbitals from subsystem B (the environment) that capture the important environment relaxation effects upon excitation. A critical distinction should be made that the orbitals chosen to add to the active region are occupied MOs from the environment, not the virtual orbitals. The reason for this is that the embedding method does not restrict the virtuals of the environment, and thus they are free to mix with the active region orbitals to obtain accurate correlation energies as they would in a traditional calculation.

As an example, let us describe the application of projector-based embedding to a photoactive molecule we define as subsystem A in explicit solvent that we define as subsystem B shown in Figure 1. Even for such a naive partitioning of the



**Figure 1.** Left: two-electron-in-two-orbitals schematic model provided to illustrate all possible transitions in eEOM-CCSD. Case (i) is given by a green arrow: occupied orbital (active) to virtual orbital (active). Case (ii) is given by a red arrow: occupied orbital (active) to virtual orbital (environment). Case (iii) is given by a blue arrow: occupied orbital from the environment (included in the active region) to virtual orbital (active). Case (iv) is given by a gray arrow: occupied orbital from the environment (included in the active region) to virtual orbital (environment). Right: schematic representation of the transitions allowed in the projector-embedded excited-state method (same color code is used as in the left panel). The central molecule is in the active region, while the solvent molecules (in light blue) constitute the environment. If an occupied orbital of a solvent molecule plays an important role, it can be included in the active region (highlighted solvent molecule).

chemical system, projector-based embedding naturally includes contributions (i) and (ii), which correspond to all electronic excitations within the subsystem A, shown by a green arrow in Figure 1, as well as excitations from occupied orbitals in subsystem A to virtuals in subsystem B (red arrow in Figure 1). For a more accurate description, subsystem A can be expanded to encompass additional important occupied orbitals from the solvent, this would include the two further types of contribution (iii) and (iv) shown as the blue and gray

transitions. The blue arrow indicates direct contributions of those occupied orbitals in subsystem B to the electronic transition itself (solvent-to-solute transitions). The gray arrow in Figure 1 highlights how occupied orbitals of the environment respond to vertical transitions in the active (e.g., solute) region. This strategy allows for a highly flexible scheme where contributions from the environment can be tested and included if necessary and will be further discussed in details in the applications presented below. In contrast to our embedding scheme, a QM/MM formalism would not incorporate the sudden polarization (or relaxation) of the MM part induced by a transition in the QM region.<sup>41</sup> Such polarization due to a change in the chromophore density would be accounted for in a polarizable force field simulation<sup>42–44</sup> or using the effective fragment potential (EFP) method,<sup>45</sup> but a direct involvement of the environment in the transition (inclusion of (ii), (iii) and/or (iv)) would imply the full inclusion of additional solvent molecules in the QM region. As mentioned above both contributions are fully incorporated in the projector embedding described in this work as well as a DFT/HF quality ground state subsystem B potential that is not present in QM/MM. A summary of the excitations possible in a range of methods can be found in Table 1. We finally note that the response of the

**Table 1. Excitations Included by Various Techniques Used to Calculate the Solvent Effects on Excited-State Properties<sup>a</sup>**

method	excitation included
QM/MM	(i)
QM/MM(polarizable)	(i)+(iv) <sup>†</sup>
naive eEOM	(i)+(ii)
extended eEOM	(i)+(ii)+(iii)*+(iv)*
full EOM-CCSD	(i)+(ii)+(iii)+(iv)

<sup>a</sup>The italic labels can be related to Figure 1, starred (\*) contributions indicate partial inclusion (i.e., only spanning nearby solvent molecules), and (†) indicates implicit inclusion. Extended eEOM is the case where the embedding region has been expanded to include additional orbitals beyond the naive choice of only orbitals highly localized on the chromophore. (i), (ii), (iii), and (iv) are described in detail in Figure 1.

environment to an excitation in the active region could also directly be incorporated in a linear-response formalism, as recently demonstrated.<sup>46</sup>

For systems where the size of the environment becomes large (>700 basis functions), reducing the number of electrons by embedding is not necessarily sufficient to make calculations feasible due to the forth order scaling of CCSD with respect to the number virtuals. For the larger calculations, we employed basis-set truncation,<sup>32</sup> which can reduce and ultimately eliminate scaling of computational cost with respect to environment size. Standard EOM-CCSD<sup>47</sup> has computational scaling of the order  $O_{(A+B)}^2 V_{(A+B)}^4$ , where  $O$  indicates the number of occupied orbitals and  $V$  the number of virtuals, and thus the total cost depends on the number of occupied MOs and unoccupied MOs in the entire system ( $a + b$ ). With projector-based embedding this becomes  $O_{(A)}^2 V_{(A+B)}^4$  where the scaling with respect to the occupied orbitals is now frozen to the size of subsystem A. The use of basis truncation further reduces the scaling to  $O_{(A)}^2 V_{(C)}^4$ , where  $V_{(C)} < V_{(A+B)}$  and the size of C is tunable according to the density of subsystem A. Hence, subsystem B can become large without significantly affecting the computational cost.<sup>32</sup>

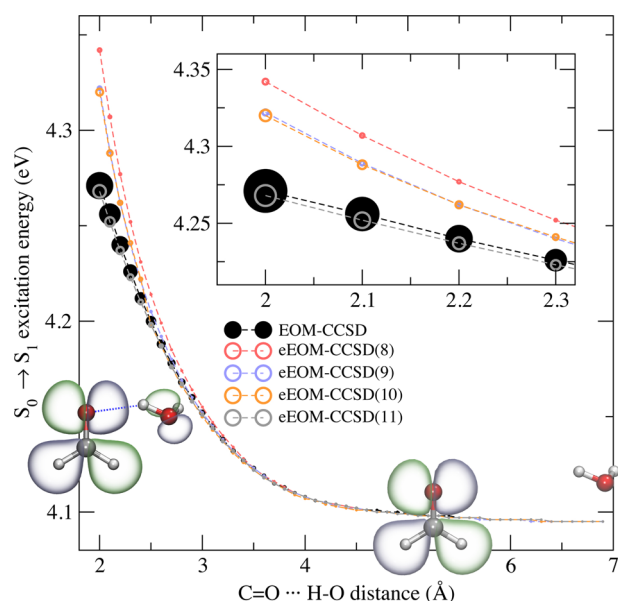
Projector embedding uses an effective embedding potential determined from the ground-state density of the full system; and the open question is whether this is appropriate for the calculation of excited-state properties, or if a full relaxation of the density in response to the excitation is necessary (for example a  $\Delta$ -SCF). In this work, we show that this is a reasonable approximation for small solvated organic molecules, provided that some of the important occupied orbitals on the solvent water molecules are included, if necessary. The ability to specify which orbitals are included in subsystem A furnishes insight into their impacts on the excitation process, and we will outline a simple and systematic approach to ensure selection of the most efficient subsystem A has been chosen.

As part of our testing, we confirmed the insensitivity of the excitation energy with the level shift parameter  $\mu$  as was done for ground state properties in the original embedding paper, indicating that the projected orbitals remain sufficiently orthogonal to the virtual space. It should also be noted that it is possible to run LR-TDDFT-in-DFT and CIS-in-HF embedding for low cost determination of the additional occupied orbitals to be included in subsystem A or merely to further accelerate these techniques.

To explore the effect of embedding potentials on excited-state properties affected by solvation, we first present results on a small model system made of a formaldehyde molecule with a single water positioned near the carbonyl group, forming an hydrogen bond. The  $S_1$  state of the isolated formaldehyde has  $1(n\pi^*)$  character, but might be perturbed by a water molecule in the close vicinity of the carbonyl group. Hence, choosing this system allows us to control the perturbation of the  $S_0 \rightarrow S_1$  excitation as well as having an easily calculable full EOM-CCSD benchmark number.

Determining the number of orbitals required to reproduce the full EOM-CCSD is critical in gauging whether projector-based embedding is feasible. If all the MOs are needed, the embedding approach would have the same cost as full EOM-CCSD and the method would have little advantage. A typical example would be an excited-state described by excitations from many orbitals, all contributing with small CI coefficients. Figure 2 shows the transition energy and oscillator strength of the  $S_0 \rightarrow S_1$  transition as a function of the distance between the nearest hydrogen on the water and the oxygen of the carbonyl group for different numbers of embedded orbitals for EOM-CCSD-in-PBE0. At the closest distance of 2 Å, we expect any error due to embedding to be most pronounced. With the minimal embedding region of 8 occupied orbitals, denoted “eEOM-CCSD(8)”, where only the orbitals on the formaldehyde are treated at the EOM-CCSD level (in red in Figure 2), the excitation energy deviates from the benchmark number in black by up to 0.7 eV. The error indicates that although the external potential is polarising the formaldehyde at the mean-field level, the relaxation of the environment necessary to construct the  $S_1$  state is still missing. The error is expected to gradually decrease by adding more occupied orbitals to subsystem A. For a separation of 2 Å, it is only when 3 additional MOs from the water are included (eEOM-CCSD(11) in Figure 2) that the energy reproduces the full EOM answer. This observation can be rationalized by looking at the HOMO orbital plots (Figure 2) which show a significant contribution of the water orbital at short water-formaldehyde distance. Such contributions disappear at larger distances. As the water distance increases the contribution of the water occupied orbitals to the excitation becomes smaller, and at 3 Å

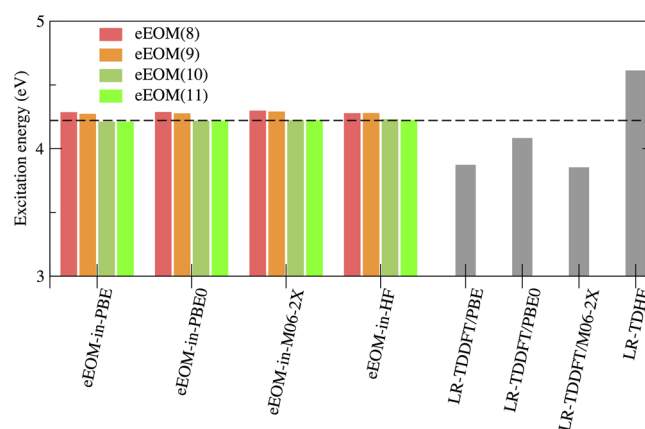




**Figure 2.**  $S_0 \rightarrow S_1$  transition energy for the formaldehyde+water system at the eOM-CCSD and EOM-CCSD level of theory, as a function of the  $C=O \cdots H-O$  distance. The color of the lines indicates the number of occupied orbitals treated at the eOM-CCSD level and the size of the dots indicates the oscillator strength at each  $C=O \cdots H-O$  distance. The black line is the full EOM-CCSD result which corresponds to all 13 MOs being embedded.

the difference between the number of orbitals in the embedding region becomes negligible, i.e., the occupied orbitals on the formaldehyde only (eOM-CCSD(8)) contribute to the transition. The oscillator strength (indicated by dot size in Figure 2) is near zero for this formally forbidden  $^1(n\pi^*)$  excitation and only increases when the water starts to break symmetry at shorter distance. eOM-CCSD(8) at 2 Å underestimates the oscillator strength, while eOM-CCSD(11) is within a factor of two of the benchmark value of  $6.63 \times 10^{-6}$  (inset of Figure 2). This simple test case highlights how embedding can accurately describe the excitation at a range of solute–solvent distances and also provide useful microscopic insight into the role that the solvent plays in modulating the transition energies. Comparing EOM-CCSD to eOM-CCSD results shows that at short distances, the water solvent actively participates in the transition, and additional orbitals are required to obtain full EOM-CCSD results. At longer distances, solvent orbitals are not required because it plays a more passive role, as a sort of polarizing spectator to the transition.

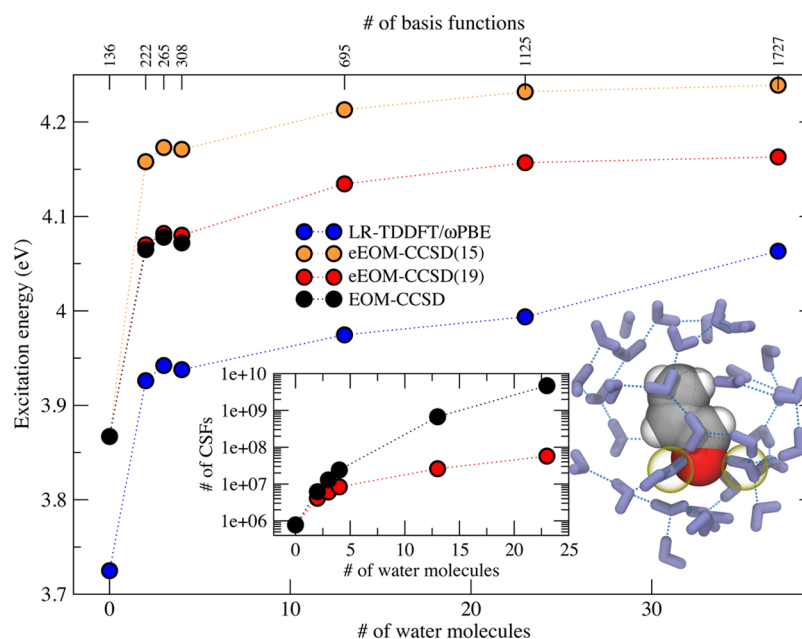
To evaluate the extent to which the choice of density functional impacts accuracy, we repeated the above test of formaldehyde+water for the 2 Å strong interaction limit and varied the functional used in the environment. We present a selection of results from four main exchange and correlation functionals that cover the families of DFT functionals generally used: PBE, PBE0, M06-2X. We also present further comparisons with HF. Additional functionals were examined and we observed similar trends as what is presented in the following. From Figure 3 it can be seen that, even for this strongly interacting limit, the error in excitation energy is 0.07 eV for the 8 and 9 orbital case and for the 10 and 11 orbital case is within 0.02 eV. Hence, for localized transitions similar to the one under study here, there is only a weak dependence on



**Figure 3.** Effect of the environment potential on the  $S_0 \rightarrow S_1$  transition energy on the formaldehyde+water for eOM-in-DFT and eOM-in-HF (left). The LR-TDDFT (with the same set of functionals) and LR-TDHF result is reported in gray on the right. The black dashed line is the reference all-electron EOM-CCSD. All calculations used an aug-cc-pVTZ basis set.

the choice of density functional approximation. Such weak dependence is clearly not observed with LR-TDDFT (and TDHF) when using the same set of functionals, as shown in the right part of Figure 3. These results further show that the most important factor in the embedding is the selection of environment orbitals in the active region to be treated at the high-level. To test the role of the orbital selection, we investigated strategies beyond the default Mulliken scheme. Specifically, we identified orbitals which had the largest energetic impact, and selected those. Using this strategy for example, the error of embedding 9 orbitals was reduced from 0.067 to 0.022 eV for the M06-2X functional. This improvement indicates that orbital choices based on Mulliken population analysis are not optimal for finding excited-state properties and that a better criterion is possible (likely involving analysis of a CIS vector).

We now move to the study of the embedded EOM-CCSD (eOM-CCSD) treatment of electronic transitions in larger molecular systems. To expand upon the work on  $^1(n\pi^*)$  transitions in formaldehyde+water to larger systems, we used a single structure from the work of Aidas<sup>43</sup> et al., which comprises an acrolein molecule solvated by a large number of water molecules, offering a realistic model of acrolein in aqueous solution. Substructures based on a radial cutoff around the acrolein center-of-mass were then selected to change the number of water molecules in the system, and an example of the largest cutoff of 5 Å is shown in Figure 4. In Figure 4, we show the results of different numbers of water molecules solvating acrolein for LR-TDDFT/ $\omega$ PBE, all electron EOM-CCSD, eOM-CCSD on the acrolein only (eOM-CCSD(15)), and eOM-CCSD with an additional 4 molecular orbitals from the 2 water molecules that are closest to the carbonyl group (eOM-CCSD(19)). An aug-cc-pVDZ basis was used for all calculations. For acrolein in gas phase (0-water case), the EOM-CCSD and eOM-CCSD(15) are the same calculation as there is no environment to embed. The  $\omega$ PBE/LR-TDDFT transition is 0.15 eV smaller than the EOM-CCSD result, a difference that is carried forward throughout the calculations. The first microsolvated system (2 water molecules next to the carbonyl group of acrolein) shows that eOM-CCSD(15), where only the occupied orbitals on acrolein are



**Figure 4.** Excitation energy of the  $S_0 \rightarrow S_1$  transition ( $^1(n\pi^*)$ ) transition in acrolein for LR-TDDFT/TDA/ $\omega$ PBE, EOM-CCSD, and embedded EOM-CCSD-in-HF on the acrolein only (eEOM-CCSD(15)) and on acrolein plus 4 selected water orbitals (eEOM-CCSD(19)). The number of waters surrounding acrolein is shown on the bottom axis, and the corresponding number of basis functions for each system is found on the top axis. The inset shows the number of configuration state functions for EOM-CCSD and the eEOM-CCSD(19) for the various numbers of waters. The molecular representation shows the system composed of an acrolein molecule surrounded by 37 water molecules (the two important water molecules discussed in the text are highlighted).

considered, is only within 0.09 eV of the full EOM-CCSD result. As a comparison, LR-TDDFT on the full molecular system (all orbitals included) deviates by 0.14 eV from the EOM-CCSD benchmark value. Upon inclusion of the four most strongly contributing water orbitals to the active region, the eEOM error reduces now to less than 0.01 eV (eEOM-CCSD(19) in Figure 4). We note that two of these four orbitals ( $\sigma(\text{O}-\text{H})$ ) were detected by the Mulliken analysis as required for the description of the embedding potential; the other two (lone pairs on the oxygen atoms) were chosen via a preliminary embedded CIS calculation and, therefore, contribute to the excitation energy directly. We determined these orbitals by incrementally increasing the size of the embedding region according to Mulliken population and calculating the CIS energy of the transition until agreement with a benchmark CIS had been reached to within 0.01 eV. This was only done for the first few model systems, as the same orbitals were found to be important to obtain an accurate excitation energy for all the cluster sizes. The first four embedded results were of clusters small enough to be compared against benchmark numbers of full EOM-CCSD, and showed an error of 0.01 eV. With 13 waters surrounding the acrolein, the calculation has already reached 700 basis functions and obtaining EOM-CCSD benchmarks was computationally intractable using our available computational resources. Thanks to its better scaling, eEOM-CCSD was possible. The eEOM-CCSD(19) trend is well within those observed for eEOM-CCSD(15) and LR-TDDFT. The favorable scaling of eEOM-CCSD(19) with respect to EOM-CCSD is demonstrated in the inset of Figure 4 by using the number of configuration state functions (CSFs) as a platform-independent measure of cost. EOM-CCSD has approximately two orders magnitude more CSFs than the eEOM-CCSD(19) result for a given number of water molecules.

The high-accuracy calculation of excited-state properties usually requires expensive correlation-based methods whose computational cost scales poorly with system size. We have shown in this work how projection-based embedding enables the calculation of accurate excitation energies and oscillator strengths in a more tractable way. By using DFT or HF as an environment potential, it is feasible to include the effect of solute contributions to excitation energies without explicitly including them at the high level. The implementation of this approach is simple: it requires specification of those chromophore atoms which receive high-level excitation treatment. If one desires further accuracy, it is possible to additionally include specific orbitals from the environment. This flexibility allows us to obtain important qualitative insight, for example allowing us to understand how specific environment orbitals participate in electronic transitions. We chose EOM-CCSD as the high-level method in this work due to its relatively black-box nature; however, the simplicity of projector-based embedding is such that extremely accurate methods like multireference configuration interaction (MRCI) can also straightforwardly be used. In future work, we intend to expand this work to cases where the subsystems are connected by covalent bonds, as well as designing a new orbital selector to automate the optimum selection of orbitals for excited-state embedding. These technologies will enable us to make progress in developing predictive models that furnish microscopic insight across a range of important systems, from photocatalysis to light harvesting.<sup>48,49</sup>

## COMPUTATIONAL DETAILS

The calculations conducted in this work were run with Molpro 2015<sup>50</sup> and Terachem<sup>51–53</sup> for LR-TDDFT (with and without the Tamm-Dancoff approximation (TDA) where indicated). The following mean-field methods were tested; PBE,<sup>54</sup> PBE0,<sup>55</sup>

M06-2X<sup>56</sup> and HF<sup>57</sup> with the aug-cc-pVDZ and aug-cc-pVTZ<sup>58</sup> basis set. For orbital localization, we utilize the intrinsic bond orbitals method of Knizia.<sup>59</sup> The basis truncation value of  $5 \times 10^{-6}$  used for the larger acrolein systems was obtained from benchmarks of smaller systems and tuned to find an error of the order of 0.01 eV. The largest calculations we performed in Figure 4 would correspond to 400 electrons in 1727 functions for a traditional EOM-CCSD.

The approach to determine which orbitals of the environment to be included was as follows: (i) HF on the whole system; (ii) select the atoms of the photoactive molecule (or part of the molecule); (iii) run CIS (or LR-TDDFT, CC2, etc.) for subsystem A; (iv) increase the number of orbitals in the embedding region by 1 and check if the difference in the excitation energy is larger than some threshold value (e.g. 0.01 eV); (v) add all those orbitals whose inclusion causes the difference in CIS (or LR-TDDFT, CC2, etc.) excitation energy to change by more than the specified threshold and rerun with eEOM-CCSD.

Molecular and orbital renderings were generated with VMD.<sup>60</sup> Orbital analysis was done with Gabedit.<sup>61</sup>

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpclett.7b02500.

Example input file for running the calculations we have presented (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

\*E-mail: [simonbennie@gmail.com](mailto:simonbennie@gmail.com).

\*E-mail: [basile.curchod@bristol.ac.uk](mailto:basile.curchod@bristol.ac.uk).

\*E-mail: [fred.manby@bristol.ac.uk](mailto:fred.manby@bristol.ac.uk).

\*E-mail: [glowacki@bristol.ac.uk](mailto:glowacki@bristol.ac.uk).

### ORCID

Basile F. E. Curchod: 0000-0002-1705-473X

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors thank Casper Steinmann for providing the acrolein structure and the members of the centre of computational chemistry for helpful discussions. S.J.B. thanks the EPSRC for funding on grant number EP/M022129/1. B.F.E.C. has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 701355 (NAMDA). D.R.G. acknowledges the Royal Society for support as a university research fellow, and also EPSRC Programme Grant EP/P021123/1.

## ■ REFERENCES

- (1) Runge, E.; Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.
- (2) Petersilka, M.; Gossmann, U. J.; Gross, E. K. U. Excitation Energies from Time-Dependent Density-Functional Theory. *Phys. Rev. Lett.* **1996**, *76*, 1212–1215.
- (3) Casida, M. E. In *Recent Advances in Density Functional Methods, Part I*; Dong, D. P., Ed.; Recent Advances in Computational Chemistry; World Scientific: Singapore, 1995; pp 155–192.
- (4) Ullrich, C. A. *Time-Dependent Density-Functional Theory*; Oxford University Press: Oxford, U.K., 2012.
- (5) Marques, M. A. L.; Maitra, N. T.; Nogueira, F. M. D. S.; Gross, E. K. U.; Rubio, A. *Fundamentals of Time-Dependent Density Functional Theory*; Springer Verlag: Berlin, 2012; Vol. 837.
- (6) Laurent, A. D.; Jacquemin, D. TD-DFT benchmarks: A review. *Int. J. Quantum Chem.* **2013**, *113*, 2019–2039.
- (7) Casida, M.; Huix-Rotllant, M. Progress in Time-Dependent Density-Functional Theory. *Annu. Rev. Phys. Chem.* **2012**, *63*, 287–323.
- (8) Casida, M. E. Time-dependent density-functional theory for molecules and molecular solids. *J. Mol. Struct.: THEOCHEM* **2009**, *914*, 3–18.
- (9) Adamo, C.; Jacquemin, D. The calculations of excited-state properties with Time-Dependent Density Functional Theory. *Chem. Soc. Rev.* **2013**, *42*, 845–856.
- (10) Dreuw, A.; Head-Gordon, M. Failure of Time-Dependent Density Functional Theory for Long-Range Charge-Transfer Excited States: The Zincbacteriochlorin-Bacteriochlorin and Bacteriochlorophyll-Spheroidene Complexes. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.
- (11) Maitra, N. T. Charge-Transfer in Time-Dependent Density Functional Theory. *J. Phys.: Condens. Matter* **2017**, *29*, 423001.
- (12) Maitra, N. T.; Zhang, F.; Cave, R. J.; Burke, K. Double excitations within time-dependent density functional theory linear response. *J. Chem. Phys.* **2004**, *120*, 5932.
- (13) Elliott, P.; Goldson, S.; Canahui, C.; Maitra, N. T. Perspectives on double-excitations in TDDFT. *Chem. Phys.* **2011**, *391*, 110–119.
- (14) Levine, B. G.; Ko, C.; Quenneville, J.; Martinez, T. J. Conical intersections and double excitations in density functional theory. *Mol. Phys.* **2006**, *104*, 1039.
- (15) Stanton, J. F.; Bartlett, R. J. The equation of motion coupled cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* **1993**, *98*, 7029–7039.
- (16) Krylov, A. I. Equation-of-motion coupled-cluster methods for open-shell and electronically excited species: The hitchhiker's guide to Fock space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433.
- (17) Korona, T.; Werner, H.-J. Local treatment of electron excitations in the EOM-CCSD method. *J. Chem. Phys.* **2003**, *118*, 3006–3019.
- (18) Baudin, P.; Kristensen, K. LoFEx A local framework for calculating excitation energies: Illustrations using RI-CC2 linear response theory. *J. Chem. Phys.* **2016**, *144*, 224106.
- (19) Hofener, S.; Klopper, W. Natural transition orbitals for the calculation of correlation and excitation energies. *Chem. Phys. Lett.* **2017**, *679*, 52–59.
- (20) Dutta, A. K.; Neese, F.; Izsak, R. Towards a pair natural orbital coupled cluster method for excited states. *J. Chem. Phys.* **2016**, *145*, 034102.
- (21) Baudin, P.; Bykov, D.; Liakh, D.; Ettenhuber, P.; Kristensen, K. A local framework for calculating coupled cluster singles and doubles excitation energies (LoFEx-CCSD). *Mol. Phys.* **2017**, *115*, 2135–2144.
- (22) Kaliman, I. A.; Krylov, A. I. New algorithm for tensor contractions on multi-core CPUs, GPUs, and accelerators enables CCSD and EOM-CCSD calculations with over 1000 basis functions on a single compute node. *J. Comput. Chem.* **2017**, *38*, 842–853.
- (23) Epifanovsky, E.; Zuev, D.; Feng, X.; Khistyayev, K.; Shao, Y.; Krylov, A. I. General implementation of the resolution-of-the-identity and Cholesky representations of electron repulsion integrals within coupled-cluster and equation-of-motion methods: Theory and benchmarks. *J. Chem. Phys.* **2013**, *139*, 134105.
- (24) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.
- (25) Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.
- (26) Sneskov, K.; Schwabe, T.; Christiansen, O.; Kongsted, J. Scrutinizing the effects of polarization in QM/MM excited state calculations. *Phys. Chem. Chem. Phys.* **2011**, *13*, 18551–18560.



- (27) Sneskov, K.; Schwabe, T.; Kongsted, J.; Christiansen, O. The polarizable embedding coupled cluster method. *J. Chem. Phys.* **2011**, *134*, 104108.
- (28) Hofener, S.; Visscher, L. Wave Function Frozen-Density Embedding: Coupled Excitations. *J. Chem. Theory Comput.* **2016**, *12*, 549–557.
- (29) Daday, C.; König, C.; Neugebauer, J.; Filippi, C. Wavefunction in Density Functional Theory Embedding for Excited States: Which Wavefunctions, which Densities? *ChemPhysChem* **2014**, *15*, 3205–3217.
- (30) Olsen, J. M. H.; Steinmann, C.; Ruud, K.; Kongsted, J. Polarizable Density Embedding: A New QM/QM/MM-Based Computational Strategy. *J. Phys. Chem. A* **2015**, *119*, 5344–5355.
- (31) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. A Simple, Exact Density-Functional-Theory Embedding Scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.
- (32) Bennie, S. J.; van der Kamp, M. W.; Penniford, R. C. R.; Stella, M.; Manby, F. R.; Mulholland, A. J. A Projector-Embedding Approach for Multiscale Coupled-Cluster Calculations Applied to Citrate Synthase. *J. Chem. Theory Comput.* **2016**, *12*, 2689–2697.
- (33) Bennie, S. J.; Stella, M.; Miller, T. F.; Manby, F. R. Accelerating wavefunction in density-functional-theory embedding by truncating the active basis set. *J. Chem. Phys.* **2015**, *143*, 024105.
- (34) Goodpaster, J. D.; Barnes, T. A.; Manby, F. R.; Miller, T. F. Accurate and systematically improvable density functional theory embedding for correlated wavefunctions. *J. Chem. Phys.* **2014**, *140*, 18A507.
- (35) Isborn, C. M.; Mar, B. D.; Curchod, B. F. E.; Tavernelli, I.; Martinez, T. J. The Charge Transfer Problem in Density Functional Theory Calculations of Aqueously Solvated Molecules. *J. Phys. Chem. B* **2013**, *117*, 12189–12201.
- (36) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140*, A1133–A1138.
- (37) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618–622.
- (38) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*; Cambridge Molecular Science; Cambridge University Press: Cambridge, U.K., 2009.
- (39) Casida, M. E.; Wesolowski, T. A. Generalization of the Kohn-Sham equations with constrained electron density formalism and its time-dependent response theory formulation. *Int. J. Quantum Chem.* **2004**, *96*, 577–588.
- (40) Neugebauer, J.; Louwerse, M. J.; Baerends, E. J.; Wesolowski, T. A. The merits of the frozen-density embedding scheme to model solvatochromic shifts. *J. Chem. Phys.* **2005**, *122*, 094115.
- (41) Lin, H.; Truhlar, D. G. QM/MM: what have we learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2007**, *117*, 185.
- (42) Sneskov, K.; Matito, E.; Kongsted, J.; Christiansen, O. Approximate inclusion of triple excitations in combined coupled cluster/molecular mechanics: calculations of electronic excitation energies in solution for acrolein, water, formamide, and N-methylacetamide. *J. Chem. Theory Comput.* **2010**, *6*, 839–850.
- (43) Aidas, K.; Mogelhof, A.; Nilsson, E. J. K.; Johnson, M. S.; Mikkelsen, K. V.; Christiansen, O.; Söderhjelm, P.; Kongsted, J. On the performance of quantum chemical methods to predict solvatochromic effects: The case of acrolein in aqueous solution. *J. Chem. Phys.* **2008**, *128*, 194503.
- (44) Loco, D.; Polack, É.; Caprasecca, S.; Lagardere, L.; Lipparini, F.; Piquemal, J.-P.; Mennucci, B. A QM/MM approach using the AMOEBA polarizable embedding: from ground state energies to electronic excitations. *J. Chem. Theory Comput.* **2016**, *12*, 3654–3661.
- (45) DeFusco, A.; Minezawa, N.; Slipchenko, L. V.; Zahariev, F.; Gordon, M. S. Modeling solvent effects on electronic excited states. *J. Phys. Chem. Lett.* **2011**, *2*, 2184–2192.
- (46) Ding, F.; Tsuchiya, T.; Manby, F. R.; Miller, T. F. Linear-response time-dependent embedded mean-field theory. *J. Chem. Theory Comput.* **2017**, *13*, 4216–4227.
- (47) Sneskov, K.; Christiansen, O. Excited state coupled cluster methods. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2012**, *2*, 566–584.
- (48) Sisto, A.; Glowacki, D. R.; Martinez, T. J. Ab Initio Nonadiabatic Dynamics of Multichromophore Complexes: A Scalable Graphical-Processing-Unit-Accelerated Exciton Framework. *Acc. Chem. Res.* **2014**, *47*, 2857.
- (49) Sisto, A.; Stross, C.; van der Kamp, M. W.; O'Connor, M.; McIntosh-Smith, S.; Johnson, G. T.; Hohenstein, E. G.; Manby, F. R.; Glowacki, D. R.; Martinez, T. J. Atomistic non-adiabatic dynamics of the LH2 complex with a GPU-accelerated ab initio exciton model. *Phys. Chem. Chem. Phys.* **2017**, *19*, 14924.
- (50) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; GYörfy, W.; Kats, D.; Korona, T.; Lindh, R. et al. MOLPRO, version 2015.1, a package of ab initio programs. 2015; see <http://www.molpro.net>.
- (51) Ufimtsev, I. S.; Martinez, T. J. Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *J. Chem. Theory Comput.* **2009**, *5*, 2619–2628.
- (52) Titov, A. V.; Ufimtsev, I. S.; Luehr, N.; Martinez, T. J. Generating Efficient Quantum Chemistry Codes for Novel Architectures. *J. Chem. Theory Comput.* **2013**, *9*, 213–221.
- (53) Isborn, C. M.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. Excited-State Electronic Structure with Configuration Interaction Singles and TammDancoff Time-Dependent Density Functional Theory on Graphical Processing Units. *J. Chem. Theory Comput.* **2011**, *7*, 1814–1823.
- (54) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (55) Adamo, C.; Barone, V. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* **1999**, *110*, 6158–6170.
- (56) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
- (57) Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, *23*, 69–89.
- (58) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (59) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.
- (60) Humphrey, W.; Dalke, A.; Schulten, K. VMD - Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (61) Allouche, A. Gabedit A graphical user interface for computational chemistry softwares. *J. Comput. Chem.* **2011**, *32*, 174–182.